

RAW SEQUENCE LISTING

EFS

The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number: 10/560,957
Source: 1.FWP
Date Processed by STIC: 4/5/07

ENTERED



IFWO

RAW SEQUENCE LISTING

DATE: 04/05/2007

PATENT APPLICATION: US/10/560,957

TIME: 15:52:47

Input Set : N:\EFS\04_05_07\10560957_efs\564462009500Seqlist.txt

Output Set: N:\CRF4\04052007\J560957.raw

4 <110> APPLICANT: STEER, Brian
 5 CALLEN, Walter
 6 HEALEY, Shaun
 7 PULLIAM, Derrick
 9 <120> TITLE OF INVENTION: GLUCANASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
 10 FOR MAKING AND USING THEM
 12 <130> FILE REFERENCE: 564462009500
 C--> 14 <140> CURRENT APPLICATION NUMBER: US/10/560,957
 C--> 15 <141> CURRENT FILING DATE: 2005-12-14
 17 <150> PRIOR APPLICATION NUMBER: PCT/US2004/021492
 18 <151> PRIOR FILING DATE: 2004-07-02
 20 <150> PRIOR APPLICATION NUMBER: 60/484,725
 21 <151> PRIOR FILING DATE: 2003-07-02
 23 <160> NUMBER OF SEQ ID NOS: 518
 25 <170> SOFTWARE: Patent In version 3.1
 27 <210> SEQ ID NO: 1
 28 <211> LENGTH: 1035
 29 <212> TYPE: DNA
 30 <213> ORGANISM: Unknown
 32 <220> FEATURE:
 33 <223> OTHER INFORMATION: Obtained from an environmental sample.
 35 <400> SEQUENCE: 1
 36 atgagagata taagtccctgc agagctggtt gccgagatga caaccggatg gaatcttgga 60
 37 aatacctttg atgcatatgg aaaaggcggg cttgatgatg agacaggctg gggaaatccc 120
 38 tatactacta aggaaatgat tgatgtagtc tgtgaaaagg ggtttaattc tatcagaatc 180
 39 ccaataacct gggctgatca tatgggtgct gctcctgact atacagtaga tgaggactgg 240
 40 atgaaccgtg tagaagaggt tgtaaattat gctcttgatg acgggatgta tgtcattatc 300
 41 aattcccacc acgaagagtc ctggagaatc cctgatgatg cacacattga tgcagtagat 360
 42 gaacagggtg gaaagctctg ggtccagata gctgagaggt tcagggatta tggcgaccat 420
 43 cttatttttg aggggctaaa tgagccacgt gttaagggcg gtgaaaatga gtggaatggc 480
 44 ggaacgaccg aaggacgtaa atgcctggac agacttaatc agacttttgt agattcagta 540
 45 agatcaacag gtggaaataa tgaaaaaaga cttgtactta taacaagctt tgcctcctca 600
 46 cacgtaatac agacaatagg aagccttaaa attccaagcg acgatcacct tgcctgttca 660
 47 atccatgcct atacgcctta tgattttaca tatgcctcgc gcacctctc tgagctttta 720
 48 acctgggatg gttccagaaa aagtgatatt gcttctgtta ttggtgatgt aaaaagaatc 780
 49 tttatagaca aggggtattcc tgctcctatg acagaatatg gtgcagttga taaagatggc 840
 50 aactccggtg atgtaagcgc ctgggtaact gagtatttaa cacgcgcaaa aaaagccggt 900
 51 atcccatgct tttggtggga caatggcctg tatgaatcag gtgatgaaca ttttgcata 960
 52 ttcaaccgca atgacctgac ctggtacaga gaagacgtcg ttgatgccat tatggctgtc 1020
 53 tactatgccc aataa 1035
 55 <210> SEQ ID NO: 2
 56 <211> LENGTH: 344
 57 <212> TYPE: PRT

see p. 6

RAW SEQUENCE LISTING

DATE: 04/05/2007

PATENT APPLICATION: US/10/560,957

TIME: 15:52:47

Input Set : N:\EFS\04_05_07\10560957_efs\564462009500Seqlist.txt

Output Set: N:\CRF4\04052007\J560957.raw

```

58 <213> ORGANISM: Unknown
60 <220> FEATURE:
61 <223> OTHER INFORMATION: Obtained from an environmental sample.
63 <400> SEQUENCE: 2
64 Met Arg Asp Ile Ser Pro Ala Glu Leu Val Ala Glu Met Thr Thr Gly
65 1 5 10 15
66 Trp Asn Leu Gly Asn Thr Phe Asp Ala Tyr Gly Lys Gly Gly Leu Asp
67 20 25 30
68 Asp Glu Thr Gly Trp Gly Asn Pro Tyr Thr Thr Lys Glu Met Ile Asp
69 35 40 45
70 Val Val Cys Glu Lys Gly Phe Asn Ser Ile Arg Ile Pro Ile Thr Trp
71 50 55 60
72 Ala Asp His Met Gly Ala Ala Pro Asp Tyr Thr Val Asp Glu Asp Trp
73 65 70 75 80
74 Met Asn Arg Val Glu Glu Val Val Asn Tyr Ala Leu Asp Asp Gly Met
75 85 90 95
76 Tyr Val Ile Ile Asn Ser His His Glu Glu Ser Trp Arg Ile Pro Asp
77 100 105 110
78 Asp Ala His Ile Asp Ala Val Asp Glu Gln Val Gly Lys Leu Trp Val
79 115 120 125
80 Gln Ile Ala Glu Arg Phe Arg Asp Tyr Gly Asp His Leu Ile Phe Glu
81 130 135 140
82 Gly Leu Asn Glu Pro Arg Val Lys Gly Gly Glu Asn Glu Trp Asn Gly
83 145 150 155 160
84 Gly Thr Thr Glu Gly Arg Lys Cys Leu Asp Arg Leu Asn Gln Thr Phe
85 165 170 175
86 Val Asp Ser Val Arg Ser Thr Gly Gly Asn Asn Glu Lys Arg Leu Val
87 180 185 190
88 Leu Ile Thr Ser Phe Ala Ser Ser His Val Ile Gln Thr Ile Gly Ser
89 195 200 205
90 Leu Lys Ile Pro Ser Asp Asp His Leu Ala Val Ser Ile His Ala Tyr
91 210 215 220
92 Thr Pro Tyr Asp Phe Thr Tyr Ala Ser Gly Thr Ser Ser Glu Leu Leu
93 225 230 235 240
94 Thr Trp Asp Gly Ser Arg Lys Ser Asp Ile Ala Ser Val Ile Gly Asp
95 245 250 255
96 Val Lys Arg Ile Phe Ile Asp Lys Gly Ile Pro Val Leu Met Thr Glu
97 260 265 270
98 Tyr Gly Ala Val Asp Lys Asp Gly Asn Ser Gly Asp Val Ser Ala Trp
99 275 280 285
100 Val Thr Glu Tyr Leu Thr Arg Ala Lys Lys Ala Gly Ile Pro Cys Phe
101 290 295 300
102 Trp Trp Asp Asn Gly Leu Tyr Glu Ser Gly Asp Glu His Phe Ala Ile
103 305 310 315 320
104 Phe Asn Arg Asn Asp Leu Thr Trp Tyr Arg Glu Asp Val Val Asp Ala
105 325 330 335
106 Ile Met Ala Val Tyr Tyr Ala Gln
107 340
109 <210> SEQ ID NO: 3

```

RAW SEQUENCE LISTING

DATE: 04/05/2007

PATENT APPLICATION: US/10/560,957

TIME: 15:52:47

Input Set : N:\EFS\04_05_07\10560957_efs\564462009500Seqlist.txt

Output Set: N:\CRF4\04052007\J560957.raw

```

110 <211> LENGTH: 1122
111 <212> TYPE: DNA
112 <213> ORGANISM: Unknown
114 <220> FEATURE:
115 <223> OTHER INFORMATION: Obtained from an environmental sample.
117 <400> SEQUENCE: 3
118 gctgacctgc ggcgacgtcg cctgctgcag ggggcccga cgtgcccgt gcttggtggtg 60
119 tgcagtgcgc aggcgcgcgc cgcaccgcgc tggccggcgt ggcaggtaact gctcgacagc 120
120 agcctgagcc ggcgacggacg catgatcgat cgcagccagg acgatcagcg cagcacctcc 180
121 gaaggccagt cctatgcgtt gttcttcgcc ctggtcgaca acgaccaggc gctgttcgac 240
122 cgcacctctgg gctggaccca ggacaacctc gccgggtggcg acatgcgcca gcacctgccc 300
123 gcctggttgt gggggcggga tgacaaggga agctggcggg tactggacga caaccggcc 360
124 tccgactccg atctctggct ggcctatgcg ctgctggaag gcgcgaggct gtggcgccgt 420
125 cccgcgttga aggccatcgc cgagggcctg ctgcgcagg tgcgggcacg cgagatcgtg 480
126 gacctgcccg ggctcggccc gatgctgttg cggggcccac agggattcac cgaagggggac 540
127 gcgacgcggg tcaaccccag ctatctgccca ctgccgtgc tgcgcgctt tgcggtggag 600
128 gatcgacgag ggcctgtggc ggcgtggcg cgtaacagcg tgcagctgtt gcagcagacc 660
129 agcccaaagg ggttcgcacc ggactggcg gcgtggaaag gcgatcgttt cgtcgtcgat 720
130 ccggtgcgcg gcgcggtcgg cagctacgat gcgatccgt gctacacctg ggcaggcatg 780
131 accgcgccgc gcgatgcgtt gttccgcacg cagctggcgg cctgtccgg gccgctgcag 840
132 cgcctgcgca gtggtgcgcc gatgtgggaa aaaatcgaca cccgcagtgg ccaggggcag 900
133 ggcaaggca actacggatt ccgcgcggca ctgctgccgt atctgatcgc gcagggcgac 960
134 gcggagcgcg cgcagtcgct gcgggcccag ctgccagcg ccgaacagca gcgtgctgat 1020
135 gcaccggcct actactcgca gatgctggcc ctgttcggcc tgggctgggc cgaagggcg 1080
136 tggcgtttcg ccgccgacgg ccgcctgcag ccgcgtggt ag 1122
138 <210> SEQ ID NO: 4
139 <211> LENGTH: 373
140 <212> TYPE: PRT
141 <213> ORGANISM: Unknown
143 <220> FEATURE:
144 <223> OTHER INFORMATION: Obtained from an environmental sample.
146 <220> FEATURE:
147 <221> NAME/KEY: SIGNAL
148 <222> LOCATION: (1)...(25)
150 <400> SEQUENCE: 4
151 Ala Asp Leu Arg Arg Arg Arg Leu Leu Gln Ala Ala Ala Thr Leu Pro
152 1 5 10 15
153 Leu Leu Gly Trp Cys Ser Ala Gln Ala Ala Pro Ala Pro Arg Trp Pro
154 20 25 30
155 Ala Trp Gln Val Leu Leu Asp Ser Ser Leu Ser Arg Asp Gly Arg Met
156 35 40 45
157 Ile Asp Arg Ser Gln Asp Asp Gln Arg Ser Thr Ser Glu Gly Gln Ser
158 50 55 60
159 Tyr Ala Leu Phe Phe Ala Leu Val Asp Asn Asp Gln Ala Leu Phe Asp
160 65 70 75 80
161 Arg Ile Leu Gly Trp Thr Gln Asp Asn Leu Ala Gly Gly Asp Met Arg
162 85 90 95
163 Gln His Leu Pro Ala Trp Leu Trp Gly Arg Asp Asp Lys Gly Ser Trp
164 100 105 110

```

RAW SEQUENCE LISTING

DATE: 04/05/2007

PATENT APPLICATION: US/10/560,957

TIME: 15:52:47

Input Set : N:\EFS\04_05_07\10560957_efs\564462009500Seqlist.txt

Output Set: N:\CRF4\04052007\J560957.raw

```

165 Arg Val Leu Asp Asp Asn Pro Ala Ser Asp Ser Asp Leu Trp Leu Ala
166      115      120      125
167 Tyr Ala Leu Leu Glu Gly Ala Arg Leu Trp Arg Arg Pro Ala Leu Lys
168      130      135      140
169 Ala Ile Ala Glu Gly Leu Leu Ala Gln Val Arg Ala Arg Glu Ile Val
170 145      150      155      160
171 Asp Leu Pro Gly Leu Gly Pro Met Leu Leu Pro Gly Pro Gln Gly Phe
172      165      170      175
173 Thr Glu Gly Asp Ala Thr Arg Val Asn Pro Ser Tyr Leu Pro Leu Pro
174      180      185      190
175 Leu Leu Arg Arg Phe Ala Val Glu Asp Arg Ser Gly Pro Trp Gln Ala
176      195      200      205
177 Leu Ala Arg Asn Ser Val Gln Leu Leu Gln Gln Thr Ser Pro Lys Gly
178      210      215      220
179 Phe Ala Pro Asp Trp Ala Ala Trp Lys Gly Asp Arg Phe Val Val Asp
180 225      230      235      240
181 Pro Val Arg Gly Ala Val Gly Ser Tyr Asp Ala Ile Arg Cys Tyr Thr
182      245      250      255
183 Trp Ala Gly Met Thr Ala Pro Arg Asp Ala Leu Phe Arg Thr Gln Leu
184      260      265      270
185 Ala Ala Leu Ser Gly Pro Leu Gln Arg Leu Arg Ser Gly Ala Pro Met
186      275      280      285
187 Trp Glu Lys Ile Asp Thr Arg Ser Gly Gln Gly Gln Gly Glu Gly Asn
188      290      295      300
189 Tyr Gly Phe Arg Ala Ala Leu Leu Pro Tyr Leu Ile Ala Gln Gly Asp
190 305      310      315      320
191 Ala Glu Arg Ala Gln Ser Leu Arg Ala Ser Leu Pro Ser Ala Glu Gln
192      325      330      335
193 Gln Arg Ala Asp Ala Pro Ala Tyr Tyr Ser Gln Met Leu Ala Leu Phe
194      340      345      350
195 Gly Leu Gly Trp Ala Glu Gly Arg Trp Arg Phe Ala Ala Asp Gly Arg
196      355      360      365
197 Leu Gln Pro Arg Trp
198      370

```

201 <210> SEQ ID NO: 5

202 <211> LENGTH: 954

203 <212> TYPE: DNA

204 <213> ORGANISM: Bacteria

206 <400> SEQUENCE: 5

```

207 atgggtgttg atccttttga aaggaacaaa atattgggaa gaggcattaa tataggaaat      60
209 gcgcttgaag caccaaatga gggagactgg ggagtgggtga taaaagatga gttcttcgac      120
211 attataaaag aagccggttt ctctcatggt cgaattccaa taagatggag tacgcacgct      180
213 tacgcgtttc ctctttataa aatcatggat cgcttcttca aaagagtgga tgaagtgata      240
215 aacggagccc tgaaaagagg actggctggt gttataaata ttcatacta cgaggagtta      300
217 atgaatgata cagaagaaca caaggaaaga tttcttgctc tttggaaaca aattgctgat      360
219 cgttataaag actatcccga aactctat ttgaaattc tgaatgaacc tcacggaaat      420
221 cttactccgg aaaaatggaa tgaactgctt gaggaagctc taaaagttaa aagatcaatt      480
223 gacaaaaagc acactataat tataggcaca gctgaatggg ggggtatatc tgcccttgaa      540
225 aaactgtctg tcccaaaatg ggaaaaaaat tctatagtta caattcacta ctacaatcct      600

```

RAW SEQUENCE LISTING

DATE: 04/05/2007

PATENT APPLICATION: US/10/560,957

TIME: 15:52:47

Input Set : N:\EFS\04_05_07\10560957_efs\564462009500Seqlist.txt

Output Set: N:\CRF4\04052007\J560957.raw

```

227 ttcgaattta cccatcaagg agctgagtgg gtggaaggat ctgagaaatg gttgggaaga      660
229 aagtggggat ctccagatga tcagaaacat ttgatagaag aattcaattt tatagaagaa      720
231 tggtaaaaaa agaacaaaag accaatttac ataggtgagt ttggtgccta cagaaaagct      780
233 gaccttgaat caagaataaa atggacctcc tttgtcgttc gcgaaatgga gaaaaggaga      840
235 tggagctggg catactggga attttgttcc ggttttggtg tttatgatac tctgagaaaa      900
237 acctggaata aagatctttt agaagcttta ataggaggag atagcattga ataa          954
240 <210> SEQ ID NO: 6
241 <211> LENGTH: 317
242 <212> TYPE: PRT
243 <213> ORGANISM: Bacteria
245 <220> FEATURE:
246 <221> NAME/KEY: DOMAIN
247 <222> LOCATION: (19)...(296)
248 <223> OTHER INFORMATION: Cellulase (glycosyl hydrolase family 5)
250 <400> SEQUENCE: 6
251 Met Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg Gly Ile
252 1 5 10 15
255 Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp Gly Val
256 20 25 30
259 Val Ile Lys Asp Glu Phe Phe Asp Ile Ile Lys Glu Ala Gly Phe Ser
260 35 40 45
263 His Val Arg Ile Pro Ile Arg Trp Ser Thr His Ala Tyr Ala Phe Pro
264 50 55 60
267 Pro Tyr Lys Ile Met Asp Arg Phe Phe Lys Arg Val Asp Glu Val Ile
268 65 70 75 80
271 Asn Gly Ala Leu Lys Arg Gly Leu Ala Val Val Ile Asn Ile His His
272 85 90 95
275 Tyr Glu Glu Leu Met Asn Asp Pro Glu Glu His Lys Glu Arg Phe Leu
276 100 105 110
279 Ala Leu Trp Lys Gln Ile Ala Asp Arg Tyr Lys Asp Tyr Pro Glu Thr
280 115 120 125
283 Leu Phe Phe Glu Ile Leu Asn Glu Pro His Gly Asn Leu Thr Pro Glu
284 130 135 140
287 Lys Trp Asn Glu Leu Leu Glu Glu Ala Leu Lys Val Ile Arg Ser Ile
288 145 150 155 160
291 Asp Lys Lys His Thr Ile Ile Ile Gly Thr Ala Glu Trp Gly Gly Ile
292 165 170 175
295 Ser Ala Leu Glu Lys Leu Ser Val Pro Lys Trp Glu Lys Asn Ser Ile
296 180 185 190
299 Val Thr Ile His Tyr Tyr Asn Pro Phe Glu Phe Thr His Gln Gly Ala
300 195 200 205
303 Glu Trp Val Glu Gly Ser Glu Lys Trp Leu Gly Arg Lys Trp Gly Ser
304 210 215 220
307 Pro Asp Asp Gln Lys His Leu Ile Glu Glu Phe Asn Phe Ile Glu Glu
308 225 230 235 240
311 Trp Ser Lys Lys Asn Lys Arg Pro Ile Tyr Ile Gly Glu Phe Gly Ala
312 245 250 255
315 Tyr Arg Lys Ala Asp Leu Glu Ser Arg Ile Lys Trp Thr Ser Phe Val
316 260 265 270

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/560,957

DATE: 04/05/2007
TIME: 15:52:48

Input Set : N:\EFS\04_05_07\10560957_efs\564462009500Seqlist.txt
Output Set: N:\CRF4\04052007\J560957.raw

FYI

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:223; N Pos. 1759,1760,1761

Seq#:224; Xaa Pos. 587

VERIFICATION SUMMARY

DATE: 04/05/2007

PATENT APPLICATION: US/10/560,957

TIME: 15:52:48

Input Set : N:\EFS\04_05_07\10560957_efs\564462009500Seqlist.txt

Output Set: N:\CRF4\04052007\J560957.raw

L:14 M:270 C: Current Application Number differs, Replaced Current Application Number

L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:11739 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:223 after pos.:1740

L:11833 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:224 after pos.:576